CRYSTALLIZATION OF RANDOM TRIGONOMETRIC POLYNOMIALS

DAVID W. FARMER AND MARK YERRINGTON

Abstract. We give a precise measure of the rate at which repeated differentiation of a random trigonometric polynomial causes the roots of the function to approach equal spacing. This can be viewed as a toy model of crystallization in one dimension. In particular we determine the asymptotics of the distribution of the roots around the crystalline configuration and find that the distribution is not Gaussian.

1. Introduction

The critical points of an analytic function have a variety of interesting physical interpretations. For polynomials we have the Gauss electrostatic model: at each zero of the polynomial $f(z)$ place identical point charges obeying an inverse linear law. Then the zeros of the derivative $f'(z)$ are the points where the field vanishes. To see when this works, just write $f(z)$ in factored form and consider the logarithmic derivative $f'(z)/f(z)$.

The Gauss model extends to entire functions of order 1, provided one incorporates a background field coming from the exponential factors of the Hadamard factorization of the function. For such functions there is a general phenomenon that differentiation smooths out irregularities in the distribution of zeros. See [6] for details. If the zeros are located in a strip around the real axis and their initial distribution is not too irregular, then repeated differentiation leads the zeros to approach equal spacing. This regular spacing is known as the *crystalline configuration* [2] and we view the process of repeated differentiation as a toy model of crystallization in one dimension. According to the Gauss model, regular spacing is the equilibrium position and differentiation moves the function toward equilibrium. See Figure 1 for an illustration.

We now describe the functions we study and then discuss our results.

A random trigonometric polynomial of degree N is a function of the form

$$F(x) = \sum_{n=0}^{N} a_n \cos(nx) + b_n \sin(nx) \tag{1.1}$$

where a_n and b_n are random variables. In this paper we will assume that the a_n and b_n are independent real Gaussian distributed with mean 0 and variance σ_n, and usually we further assume that the variances σ_n are all equal.

Our concern is with the properties of the *real* zeros of $F(x)$. Generally $F(x)$ will not have all of its zeros real, but the high derivatives of $F(x)$ will have mostly real zeros, and those zeros will be close to equally spaced. This can be seen in Figure 1. This is a general property [6] of real entire functions of order 1, but in the case of trigonometric polynomials

Research of both authors supported by the American Institute of Mathematics and the NSF Focused Research Group grant DMS 0244660.

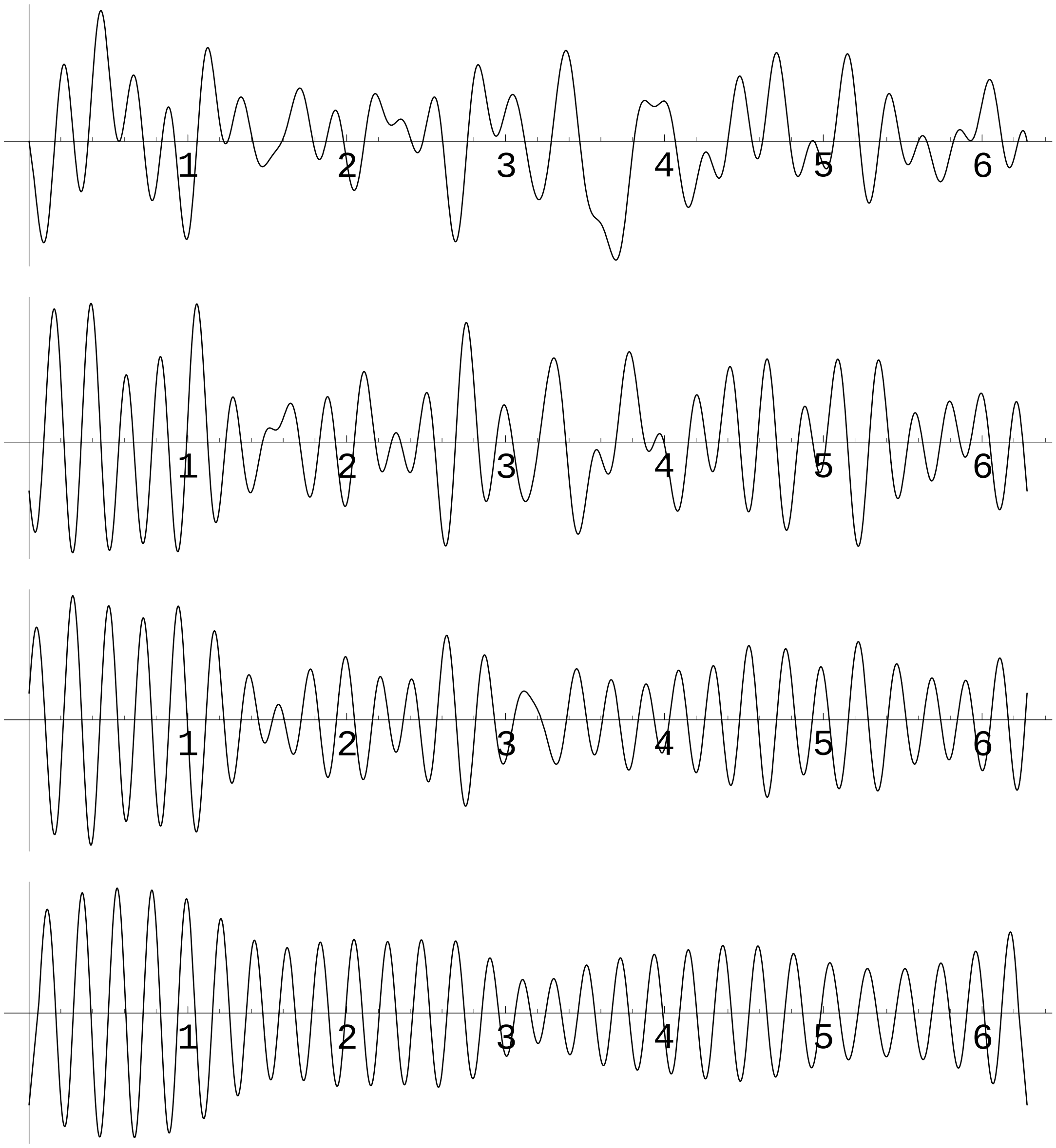

FIGURE 1. An example degree 30 trigonometric polynomial, along with its 1st, 3rd, and 10th derivative.

this is easy to see. The pth derivative of $F(x)$ is

$$F^{(p)}(x) = \sum_{n=1}^{N} a_n n^p \cos(nx) + b_n n^p \sin(nx),$$

where for simplicity we have assumed p is a multiple of 4, For large p the terms $a_N N^p \cos(Nx) + b_N N^p \sin(Nx)$ dominate. So the zeros of $F^{(p)}(x)$ are close to the zeros of $a_N \cos(Nx) + b_N \sin(Nx) = c_N \cos(Nx + \phi_N)$ for some real c_N and ϕ_N, and those zeros are real and equally spaced.

Another general property of repeated differentiation of real entire functions of order 1 is that the discrepancy from equal spacing of zeros of the pth derivative scales as $O(1/p)$. See

Theorem 2.4.2 of [6]. In the case of random trigonometric polynomials we are able to obtain more precise information. We describe this in the next section.

2. Statement of results

In this section $F(x)$ is a random trigonometric polynomial of the form (1.1) for which the a_k and b_k are independent Gaussian distributed random variables with mean 0 and identical variance. We wish to measure the rate at which the real zeros of the pth derivative $F^{(p)}(x)$ approach equal spacing as $p \to \infty$. We consider the pair correlation function of the zeros of $F^{(p)}(x)$, defined as

$$R_{2,p}(\tau) = \langle \rho_p(x)\rho_p(x+\tau)\rangle, \tag{2.1}$$

where

$$\rho_p(x) = \sum_{x_k : F^{(p)}(x_k)=0} \delta(x - x_k) = \delta(F^{(p)}(x))|F^{(p+1)}(x)|. \tag{2.2}$$

Here δ is the Dirac δ-function at 0, and $\langle\cdot\rangle$ stands for expected value. Thus, ρ_p is the density function of real zeros of $F^{(p)}$, and $R_{2,p}$ is the density function of differences of real zeros.

Since $R_{2,p}$ measures the differences between zeros, if the zeros are almost regularly spaced then $R_{2,p}(x)$ will be large when x is close to a multiple of the average zero spacing and it will be small otherwise. In other words, we expect that $R_{2,p}$ should approach a sum of δ-functions at the integers as $p \to \infty$. Bogomolny, Bohigas, and Lebœuf [2] obtain a general expression for the pair correlation function of the real zeros of a random trigonometric polynomial. From their results (which we describe in Section 4) we obtain a formula for $R_{2,p}$, which we plot for $p = 0, 1, 3, 10$ in Figure 2. Note that the $p = 0$ case is from [2].

The plots in Figure 2 use the following normalization. We rescale the polynomial so that the average spacing between zeros is 1, that is, we are actually considering the function $F^{(p)}(\pi x/N)$. As we describe in Section 3, as $N \to \infty$ the function $F^{(p)}$ has the expected fraction

$$v_p = \sqrt{\frac{2p+1}{2p+3}} \sim 1 - \frac{1}{2p} \tag{2.3}$$

of real zeros. So $1/v_p$ is the average gap between consecutive real zeros, and that is the spacing between the peaks in the pair correlation functions in Figure 2. Note that v_p is also the density of real zeros. So the pair correlation function $R_{2,p}(x)$ will equal v_p^2 on average, which can also be seen in Figure 2.

We give an asymptotic formula for the pair correlation function $R_{2,p}(x)$ as $p \to \infty$. As in the general case [6] we find a $O(1/p)$ discrepancy from equal spacing, and furthermore we find that the nearest-neighbor spacing has (appropriately rescaled) distribution function

$$\frac{1}{(1+4x^2)^{\frac{3}{2}}}, \tag{2.4}$$

centered at $1 + \frac{1}{2p}$. In particular, the discrepancy from equal spacing is not Gaussian. The precise statement is

Theorem 2.1. *As $p \to \infty$, the pair correlation function $R_{2,p}(x)$ of the real zeros of $F^{(p)}$ approaches a sum of Dirac δ-functions at the nonzero integers. The δ-function near the*

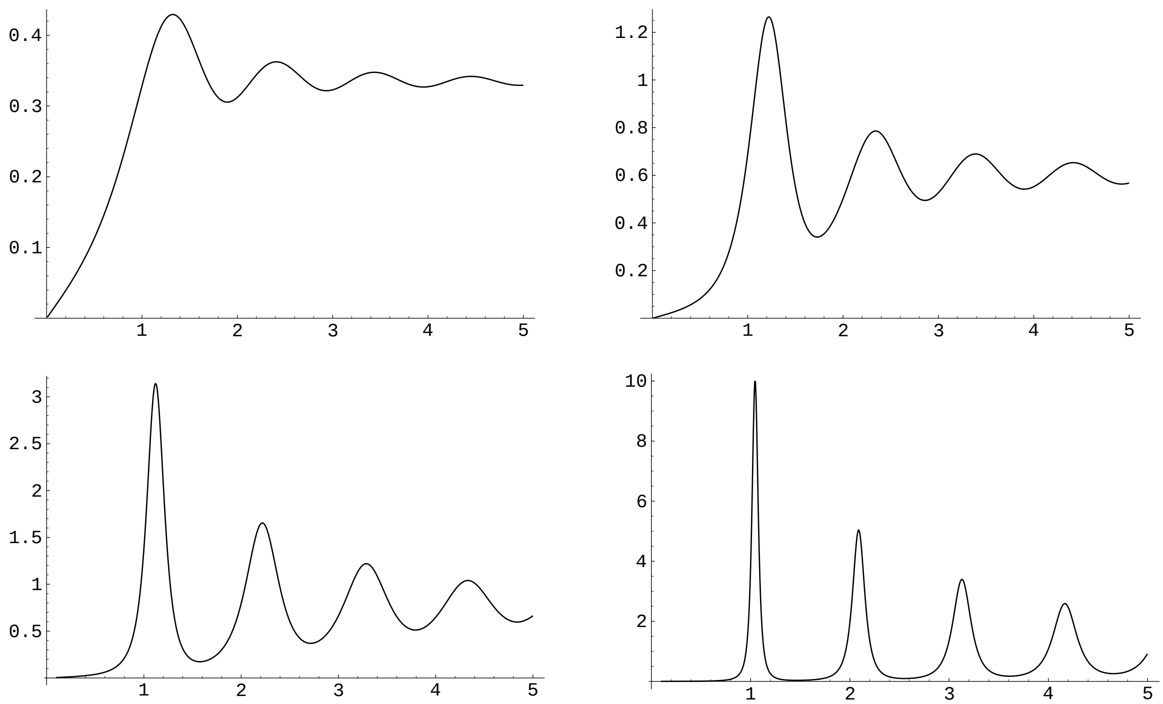

FIGURE 2. Plots of $R_{2,p}(x)$, the pair correlation funtion of the zeros of the pth derivative $F^{(p)}(x)$, for $p = 0,\ 1$ (top row) and $p = 3,\ 10$ (bottom row).

positive integer n *is given by*

$$R_{2,p}\left(n\left(1+\frac{1}{2p}+\frac{u}{p}\right)\right)=\frac{p}{n}\frac{1}{(1+4u^2)^{\frac{3}{2}}}+O(1), \tag{2.5}$$

as $p \to \infty$.

The proof is given in Section 4.1 Note that the total area under $1/(1+4u^2)^{\frac{3}{2}}$ is 1, which shows that the above Theorem does in fact identify the δ-function near the integer n.

Since the zeros of $F^{(p)}$ are close to equally spaced, the peak of $R_{2,p}(x)$ near $x = 1$ is almost completely due to nearest-neighbor spacings. Thus, we can read the nearest-neighbor distribution from the pair-correlation function, as given in (2.4). In fact, we can also read off the next-nearest neighbor spacing (and all of the other neighbor spacings), as $p \to \infty$, from the pair correlation function. Up to rescaling, all of those distributions are the same. This shows that there are long-term correlations between the zeros, otherwise, for example, the next-nearest neighbor distribution would be the convolution of the nearest neighbor distribution with itself.

One motivation for this work is to understand, in general, the effect of differentiation on the statistics of zeros for functions which have all their zeros on a line. In particular, we would like to understand the effect of differentiation on the repulsion between zeros. For example, if $f(x)$ has only real zeros and the zeros have the same statistics as the eigenvalues of the classical random matrix β-ensemble, then the zeros of $f(x)$ have repulsion of order β. For the derivative $f'(x)$, it is reasonable to conjecture that the zeros would have repulsion of order $3\beta + 1$, because pairs of close zeros of $f'(x)$ occur when $f(x)$ has three closely spaced zeros. This topic is of interest to number theory [5]. Unfortunately, the calculations in

this paper do not shed light on this phenomenon because the zeros of random trigonometric polynomials are not in general all on the real line. We find that all derivatives of a random trigonometric polynomial have *linear* repulsion between zeros, and show that this is due to the fact that each derivative moves new zeros onto the real line, and those new zeros show linear repulsion from each other. See Section 4.2.

In the next section we discuss generalities about random trigonometric polynomials. In Section 4 we discuss the asymptotics of the pair correlation of the zeros of $F^{(p)}$, and in Section 4.2 we compute the repulsion between the zeros.

3. Random trigonometric polynomials

We assume that $F(x)$ is a random trigonometric polynomial of the form (1.1), where the a_k and b_k are independent real normally distributed random variables with mean 0 and identical variance σ^2.

The expected fraction, v_p, of real zeros of $F^{(p)}(x)$ is given in (2.3). This result is due to Dunnage [3]. It follows directly from the Kac-Rice formula, which we give in Lemma 3.1.

Note that the original polynomial $F(x)$ has on average $1/\sqrt{3} \approx 57.7\%$ real zeros, the 4th derivative has more than 90% of its zeros real, and one must take 49 derivatives in order to expect 99% real zeros. Note that these are asymptotic results, and one can obtain exact formulas for any N. For example, the $N = 30$, $p = 10$ example at the bottom of Figure 1 expects to have 96.96% real zeros, and this is approximately 1.4% larger than the asymptotic estimate (2.3).

We now derive (2.3).

Lemma 3.1. *(The Kac-Rice Formula [7, 8]) Suppose $F(x)$ is a random trigonometric polynomial with independent real normally distributed coefficients having mean 0 but not necessarily equal variance, and let*

$$\begin{aligned} A^2 &= Var(F(x)) = \langle F(x)^2\rangle \\ B^2 &= Var(F'(x)) = \langle F'(x)^2\rangle \\ C &= Cov(F(x)F'(x)) = \langle F(x)F'(x)\rangle \\ \Delta^2 &= A^2B^2 - C^2. \end{aligned}$$

Then the expected number of real roots of $F(x)$ in the interval (a, b) is

$$\frac{1}{\pi}\int_a^b \frac{\Delta}{A^2}dx.$$

In the case of $F(x)$ of the form (1.1) we have

$$\begin{aligned} A^2 &= \sum_{n=0}^{N} \sigma_n^2 \\ B^2 &= \sum_{n=0}^{N} n^2\sigma_n^2 \\ C &= 0. \end{aligned}$$

If all the coefficients of $F(x)$ have equal variance σ^2, then the pth derivative $F^{(p)}(x)$ can be viewed as a random trigonometric polynomial of the form (1.1) where the coefficients a_n and b_n have variance $n^{2p}\sigma^2$. So for the pth derivative we have

$$A = \sigma^2 \sum_{n=0}^{N} n^{2p} \sim \sigma^2 \frac{N^{2p+1}}{2p+1} \tag{3.1}$$

and

$$B = \sigma^2 \sum_{n=0}^{N} n^{2p+2} \sim \sigma^2 \frac{N^{2p+3}}{2p+3}, \tag{3.2}$$

as $N \to \infty$. From this, formula (2.3) for the fraction of real roots v_p follows immediately.

4. Pair correlation of the real roots

We use a result of Bogomolny, Bohigas, and Lebœuf [2] to compute the pair correlation function $R_{2,p}$ of the real zeros of $F^{(p)}(x)$.

Lemma 4.1 (Appendix B of [2]). *Suppose*

$$F(x) = \sum_{n=0}^{N} a_n \cos(nx) + b_n \sin(nx)$$

is a random trigonometric polynomial, where the a_n and b_n are independent real Gaussian distributed random variables with mean 0 and variance σ_n. The expected value of the pair correlation function $R_2(\tau)$ of the real *zeros of $F(x)$ is given by*

$$R_2(\tau) = \frac{1}{\pi^2 C^{\frac{3}{2}}} \left(B \arcsin(B/A) + \sqrt{A^2 - B^2} \right), \tag{4.1}$$

where

$$\begin{aligned} A &= g_2 C - g_1 g_4^2 \\ B &= g_5 C - g_3 g_4^2 \\ C &= g_1^2 - g_3^2, \end{aligned} \tag{4.2}$$

with

$$\begin{aligned} g_1 &= \sum_{n=1}^{N} \sigma_n^2 \\ g_2 &= \sum_{n=1}^{N} n^2 \sigma_n^2 \\ g_3 &= \sum_{n=1}^{N} \sigma_n^2 \cos(n\tau) \\ g_4 &= \sum_{n=1}^{N} n \sigma_n^2 \sin(n\tau) \\ g_5 &= \sum_{n=1}^{N} n^2 \sigma_n^2 \cos(n\tau). \end{aligned} \tag{4.3}$$

We apply the Lemma to $F^{(p)}(x)$ with $\sigma_n = n^p$. So we have

$$g_1 \sim \frac{N^{2p+1}}{2p+1} \qquad \text{and} \qquad g_2 \sim \frac{N^{2p+3}}{2p+3}, \tag{4.4}$$

as $N \to \infty$. To determine asymptotics for g_3, g_4 and g_5, we use the fact that for continuous functions f

$$\int_0^1 f(x)dx = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} f\left(\frac{n}{N}\right).$$

We change variables $\tau = \pi x/N$, so in the variable x the mean spacing between zeros of F is unity. We have

$$\begin{aligned} g_3 &\sim N^{2p+1} \int_0^1 \cos(\pi x t)\, t^{2p} dt \\ g_4 &\sim N^{2p+2} \int_0^1 \sin(\pi x t)\, t^{2p+1} dt \\ g_5 &\sim N^{2p+3} \int_0^1 \cos(\pi x t)\, t^{2p+2} dt. \end{aligned} \tag{4.5}$$

We let $R_{2,p}(x)$ denote the pair correlation function of the real zeros of the pth derivative $F^{(p)}(x)$, where $F(x)$ is given by (1.1) with all a_j and b_j independent identical Gaussian, and we normalize by dividing by the square of the overall zero density N^2/π^2. As $N \to \infty$ we have

$$R_{2,p}(x) \sim \frac{1}{C_p^{\frac{3}{2}}} \left(B_p \arcsin(B_p/A_p) + \sqrt{A_p^2 - B_p^2} \right), \tag{4.6}$$

where

$$\begin{aligned} A_p &= g_{2,p} C_p - g_{1,p} g_{4,p}^2 \\ B_p &= g_{5,p} C_p - g_{3,p} g_{4,p}^2 \\ C_p &= g_{1,p}^2 - g_{3,p}^2, \end{aligned} \tag{4.7}$$

with

$$\begin{aligned} g_{1,p} &= \frac{1}{2p+1} \\ g_{2,p} &= \frac{1}{2p+3} \\ g_{3,p} &= \int_0^1 \cos(\pi x t)\, t^{2p} dt \\ g_{4,p} &= \int_0^1 \sin(\pi x t)\, t^{2p+1} dt \\ g_{5,p} &= \int_0^1 \cos(\pi x t)\, t^{2p+2} dt. \end{aligned} \tag{4.8}$$

Plots of $R_{2,p}$ for $p = 0, 1, 3, 10$ are given in Figure 2.

4.1. **Large p asymptotics of the pair correlation.** We determine the rate at which $R_{2,p}(x)$, appropriately rescaled, approaches a sum of δ-functions at the integers.

We first find asymptotic formulas for A_p, B_p, and C_p. Using a geometric series expansion for $g_{1,p}$ and $g_{2,p}$, and using integration by parts followed by a geometric series expansion for $g_{3,p}$, $g_{4,p}$, and $g_{5,p}$, we find (with the help of a computer algebra package), that

$$\begin{aligned} A_p &= \frac{1}{64}\left(-2\pi^2x^2 - 2\sin(2\pi x)\pi x + \left(4\pi^2x^2-1\right)\cos(2\pi x)+1\right)p^{-5} + O(p^{-6}) \\ B_p &= \frac{1}{128}\left(\cos(\pi x) + \left(4\pi^2x^2-1\right)\cos(3\pi x) - 8\pi x\sin(\pi x)\right)p^{-5} + O(p^{-6}) \\ C_p &= \frac{1}{4}\sin^2(\pi x)\,p^{-2} - \frac{1}{4}\left((\pi x\cos(\pi x)+\sin(\pi x))\sin(\pi x)\right)p^{-3} \\ &\quad + \frac{1}{32}\left(\pi^2x^2 + 8\sin(2\pi x)\pi x + 3\left(\pi^2x^2-1\right)\cos(2\pi x)+3\right)p^{-4} + O(p^{-5}) \end{aligned} \tag{4.9}$$

We see that A_p and B_p are generally of size p^{-5}, while $C_p^{3/2}$ is generally of size p^{-3}. But for $x \in \mathbb{Z}$ we see that $C_p^{3/2}$ is of size p^{-6}. Thus, as $p \to \infty$, if $x \in \mathbb{Z}$ then $R_p(x) \to \infty$, otherwise $R_p(x) \to 0$. This is exactly what one should expect because R_p is approaching a sum of δ-functions at the integers. We now express this more precisely.

The minima of C_p are not exactly at the integers, but they are shifted over to approximately $n(1+\frac{1}{2p})$ for $n \in \mathbb{Z}$. This is also what one would expect because the mean spacing of the *real* zeros of $F^{(p)}$ is approximately $1/v_p \sim 1 + \frac{1}{2p}$. We will verify this directly from the above formulas. Differentiating C_p with respect to x we have

$$C_p'(x) = -\frac{1}{16}\left(\pi\left((4p-11)\pi\cos(2\pi x)x - \pi x + \left(-4p^2+6p+3\pi^2x^2-7\right)\sin(2\pi x)\right)\right)p^{-4}. \tag{4.10}$$

We are thinking of x as fixed and p large, so the minimum of C_p is close to the solution to

$$4\pi px\cos(2\pi x) - 4p^2\sin(2\pi x) = 0, \tag{4.11}$$

which is equivalent to

$$\tan(2\pi x) = \frac{\pi x}{p}. \tag{4.12}$$

The solutions with x near an integer correspond to the minima of C_p, so writing $x = n + \xi$ and $\tan(2\pi x) \sim 2\pi\xi$, we find

$$\xi \sim \frac{n}{2p}, \qquad \text{therefore} \qquad x \sim n\left(1+\frac{1}{2p}\right), \tag{4.13}$$

as expected.

To show the shape of the δ-functions, we expand near the minima of C_p. For positive integers n we find that

$$C_p\left(n\left(1+\frac{1}{2p}+\frac{u}{p}\right)\right) = \frac{\pi^2}{16}\left(1+4u^2\right)n^2p^{-4} + O(p^{-5}). \tag{4.14}$$

Also,

$$A_p\left(n\left(1+\frac{1}{2p}+\frac{u}{p}\right)\right) = \frac{\pi^2}{32}n^2p^{-5}+O(p^{-7})$$
$$B_p\left(n\left(1+\frac{1}{2p}+\frac{u}{p}\right)\right) = \cos(\pi n)\frac{\pi^2}{32}n^2p^{-5}+O(p^{-7}). \tag{4.15}$$

Thus, $\arcsin(B_p/A_p)=(-1)^n\pi/2+O(p^{-2})$, so we obtain Theorem 2.1.

4.2. **Repulsion between zeros.** In [2] it was found the the real zeros of $F(x)$ have linear repulsion. We find that the real zeros of $F^{(p)}(x)$ also show linear repulsion.

Computing the small x asymptotics of (4.8) we find that

$$R_{2,p}(x) = \frac{\pi^2\sqrt{4p^2+8p+3}}{2(2p+3)^2(2p+5)}x+\frac{\pi^2\left(4p^2+8p+3\right)^{3/2}}{(2p+1)(2p+3)^2\sqrt{(2p+5)^3(2p+7)}}x^2+O(x^3)$$
$$\sim \frac{\pi^2}{8p^2}x, \tag{4.16}$$

as $p\to\infty$. So for every derivative we have *linear* repulsion between zeros. This appears to contradict the expectation that differentiation increases the repulsion between zeros. However, there is a simple explanation. Each derivative causes more zeros to fall onto the real line. These "new" zeros can be closely spaced, as illustrated in Figure 3.

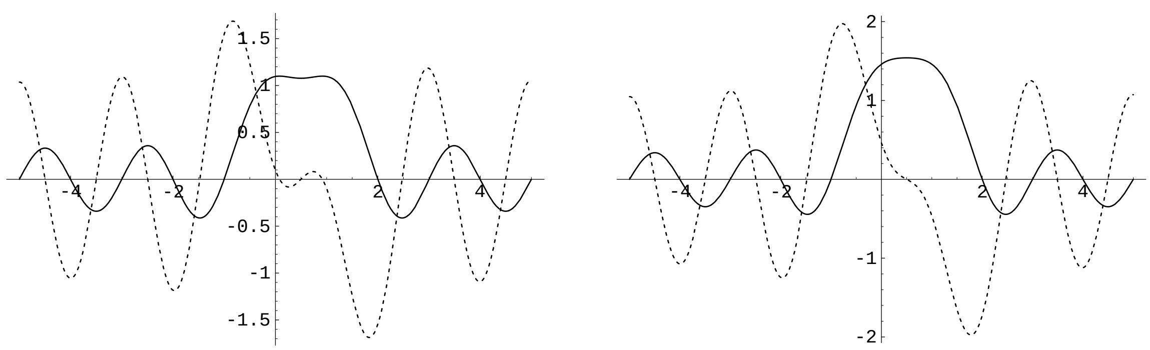

FIGURE 3. Functions with zeros at $\frac{1}{2}\pm ia$ and at the integers except for $0,1$. On the left $a=0.92$, and on the right $a=1.1$. Dotted curve shows the derivative.

In Figure 3, all of the zeros are spaced one unit apart, except for two zeros which have been moved to have imaginary part $\pm a$ and real part the midpoint of the resulting gap. In the left plot we have $a=0.92$ and in the right plot $a=1.1$. Moving the complex zeros to have imaginary part

$$\pm\frac{2}{\pi^2-8}\approx\pm 1.06975 \tag{4.17}$$

gives a triple zero of the derivative, so values slightly less than this give closely spaced zeros.

Since the fraction of new real zeros for the pth derivative is

$$v_p - v_{p-1} = \sqrt{\frac{2p+1}{2p+3}}-\sqrt{\frac{2p-1}{2p+1}}\sim\frac{1}{2p^2}, \tag{4.18}$$

which is of the same magnitude as the linear repulsion, we see that the new zeros lead to repulsion of magnitude x/p^2. It seems reasonable to believe that this accounts for all of the linear repulsion, but we have not been able to verify this.

References

1. A.T. Bharucha-Reid and M. Sambandham, *Random Polynomials*, Academic Press, 1986.
2. E. Bogomolny, O. Bohigas, and P. Lebœf, *Quantum Chaotic Dynamics and Random Polynomials* J. Statist. Phys. 85 (1996), no. 5-6, 639679.
3. J.E.A. Dunnage, *The number of real zeros of a random trigonometric polynomial*, Proc. London Math. Soc., 16:53-84, 1966.
4. A. Edelman and E. Kostlan, *How Many Zeros of a Random Polynomial are Real?* Bull. Amer. Math. Soc. 32, 1-37, 1995.
5. D.W. Farmer and S. Gonek, *The pair correlation of zeros of the derivative of the Riemann ξ-function*, in preparation.
6. D.W. Farmer and R.C. Rhoades, *Differentiation Evens Out Zero Spacings*, Trans. Amer. Math. Soc., 357 (2005), no.9, 3789–3811.
7. M. Kac, *On the average number of real roots of a random algebraic equation*, Bull. Am. Math. Soc 49, (1943), 314-320.
8. S.O. Rice, *The distribution of maxima of a random curve*, Amer. Jour. Math **61**, (1939), 409-416.

E-mail address: `farmer@aimath.org`
E-mail address: `my001m@mail.rochester.edu`